JOURNAL OF MULTIVARIATE ANALYSIS 30, 241–244 (1989)

Convergence Results for Maximum Likelihood Type Estimators in Multivariable ARMA Models II

R. DAHLHAUS

Institut für Mathematik, Gesamthochschule Essen, Essen, Federal Republic of Germany

AND

B. M. PÖTSCHER

Institut für Ökonometrie, OR und Systemtheorie, Technical University of Vienna, Vienna, Austria

Communicated by C. R. Rao

The consistency proof for the (Gaussian quasi) maximum likelihood estimator in multivariable ARMA models as given in Dunsmuir and Hannan (1976, *Adv. in Appl. Probab.* **8**, 339–364) rests on a certain property of the underlying parameter space, called B6 in their paper. It is not known whether the usual parameter spaces like the manifold $M(n)$ or the parameter spaces corresponding to echelon forms satisfy condition B6, since the argument given by Dunsmuir and Hannan to establish this fact is inconclusive. In Pötscher (1987, *J. Multivariate Anal.* **21** 29–52) it was shown how consistency can be proved without relying on B6 if the data generating process is Gaussian. In this note we show that the Gaussianity assumption can be replaced by ergodicity thus restoring Dunsmuir and Hannan's consistency proof to its full generality and extending it to parameter spaces which do not satisfy condition B6.

The first consistency proof for the (Gaussian quasi) maximum likelihood estimator of multivariable ARMA models not requiring ad hoc assumptions like, e.g., compactness of the underlying parameter space was given in the seminal paper by Dunsmuir and Hannan [2]. See also Deistler, Dunsmuir, and Hannan [1]. One feature of their consistency proof is that the proof makes use of a certain property of the parameter space. This property is called (B6) in Dunsmuir and Hannan [2] and

Received June 27, 1988; accepted August 12, 1988.

AMS 1980 subject classifications: Primary 62M10, 62M15; Secondary 93E10.

Key words and phrases: ARMA model, likelihood function, consistency, zeroes of spectral density.

0047-259X/89 $3.00

essentially requires the parameter space to be such that any spectral density corresponding to an element in the parameter space and having zeroes can be approximated in a certain sense by spectral densities which do not have zeroes but also correspond to an element in the parameter space. We note that in the case of univariate ARMA models the standard parameter spaces (i.e., the parameter spaces defined by prescribing an upper bound for the *AR* and *MA* order) can be easily shown to have property B6. However, in the multivariable case it is not known whether B6 holds for the most commonly used parameter spaces like $M(n)$ or echelon forms or not. An argument that B6 holds for these parameter spaces given in Dunsmuir and Hannan [2] turns out to be inconclusive; see the discussion in Pötscher [3]. Although the whole problem can be assumed away by allowing only ARMA processes with no zeroes in their spectral densities, it seems desirable to have a consistency proof also for the case where this absence of zeroes in the spectral densities is not assumed. One way to achieve this goal is, of course, to verify B6 for the standard parameter spaces (it turns out that this is not trivial and has not been accomplished up to now). Another line of attack is to give a consistency proof which does not rely on B6 at all. This note shows that this can be done, and hence restores the consistency result to its full generality as well as extends it to parameter spaces not satisfying B6. We note that such parameter spaces can, e.g., arise through restrictions placed on standard parameter spaces even in the univariate case. We also note that Theorem 4.4 in Pötscher [3] took a first step in this direction; in that theorem it was shown that consistency can be proved without B6, however, Gaussianity of the data generating process was assumed. This note shows that Gaussianity in that theorem can be replaced by the weaker condition of ergodicity. (We note that ergodicity is precisely what Dunsmuir and Hannan [2] used in their paper.)

It should be noted that the "old" version of the consistency proof employing condition B6 has still its merits for two reasons: first this version of the proof also works immediately for asymptotically stationary processes (e.g., ARMA processes plus a transient component), as is evident from the proofs in Dunsmuir and Hannan [2]; see also Pötscher [3]. Second it also works for estimators which are obtained from certain approximations to the (Gaussian quasi) likelihood, whereas the consistency proof not relying on B6 does not work for these estimators, see Pötscher [3, Theorem 4.3 and Remark (xii) on p. 50].

All notations are as in Pötscher [3]. To make the paper more easily accessible we recall that F below stands for the set of spectral densities (multiplied by 2π) implied by the set of s-dimensional ARMA models forming the parameter space. Furthermore $F_{\mathbb{R}}^{1,3}$ stands for the set of all rational spectral density matrices of dimension $s \times s$ which are nonsingular almost everywhere on the unit circle. The estimators $\hat{f}_{T,1}$, $\hat{f}_{T,2}$, $\hat{f}_{T,3}$ are slightly different variants of the Gaussian quasi maximum likelihood estimator (i.e., the likelihood is set up as if the data were Gaussian, but the asymptotic result does *not* use Gaussianity of the data). If $\hat{k}_{T,i}$ and $\hat{\Sigma}_{T,i}$ denote respectively the transfer function and the innovation covariance matrix corresponding to $\hat{f}_{T,i}$, then the convergence of $\hat{f}_{T,i}$ to f_0 given in the theorem below implies that $\hat{k}_{T,i}$ and $\hat{\Sigma}_{T,i}$ converge to k_0 and Σ_0 almost surely, where k_0 and Σ_0 are the true transfer function and innovation covariance matrix, respectively, cp. Pötscher [3].

THEOREM. *Let $F \subseteq F_{\mathbb{R}}^{1,3}$ be of finite degree and assume the data generating process $(y(t))$, $t \in \mathbb{N}$ (or $t \in \mathbb{Z}$), to be a strictly stationary and ergodic process with spectral density $(2\pi)^{-1} f_0$ (and hence zero mean). If $f_0 \in \bar{F}^{1,3}$ then $\hat{f}_{T,3}$ converges to f_0 almost surely. If $f_0 \in F$ then $\hat{f}_{T,1}$, $\hat{f}_{T,2}$, and $\hat{f}_{T,3}$ converge to f_0 almost surely.*

Proof. Without loss of generality assume that $y(t)$ is defined for $t \in \mathbb{Z}$. (In Pötscher [3] only the case $t \in \mathbb{N}$ was considered and $y(t)$ was set equal to zero for $t \leqslant 0$; of course, this difference is of no relevance at all.) Inspection of the proof of Theorem 4.4 in Pötscher [3] shows that it suffices to show the relation $\limsup_T T^{-1} y_T' \Gamma_T^{-1}(f_0)\, y_T \leqslant s$ almost surely where s is the dimension of $y(t)$ and y_T is the stacked vector of observations. Write f_0 as $|r|^{-2} P \Sigma P^*$, where P is a square matrix polynomial which is nonsingular for $|z|<1$, Σ is a positive definite matrix and r is a scalar polynomial having no zeroes for $|z| \leqslant 1$. Denoting by $\varepsilon(t)$ the one step prediction errors we have then $r(z)\, y(t) = P(z)\, \varepsilon(t)$. Clearly $\varepsilon(t)$ is—as a time invariant function of $y(t)$—strictly stationary and ergodic. Define $u(t) = r^{-1}(z)\, \varepsilon(t)$ which is a strictly stationary and ergodic autoregressive process. Its spectral density is given by $(2\pi)^{-1} g = (2\pi)^{-1} |r|^{-2} \Sigma$ and satisfies $g \in F_{\mathbb{R}}^{1,2,3}$. Then clearly $y_T = H u_{T+q}$ holds where $u_{T+q} = (u(1-q)', \ldots, u(T)')'$, $q = \deg P$, and H is the $T \times (T+q)$ matrix given by

$$H = \begin{bmatrix} P_q, \ldots\ldots, P_0, 0, \ldots, 0 \\ 0, P_q, \ldots\ldots, P_0, 0, \ldots, 0 \\ 0, \ldots, 0, P_q, \ldots\ldots, P_0 \end{bmatrix}.$$

Here the P_j are the coefficients of $P(z)$, i.e., $P(z) = \sum_{j=0}^{q} P_j z^j$. Then we have $T^{-1} y_T' \Gamma_T^{-1}(f_0)\, y_T = T^{-1} u_{T+q}' H' \Gamma_T^{-1}(f_0) H u_{T+q} \leqslant T^{-1} u_{T+q}' \Gamma_{T+q}^{-1}(g) u_{T+q}$ $\lambda_{\max}(\Gamma_{T+q}^{1/2}(g)\, H' \Gamma_T^{-1}(f_0)\, H \Gamma_{T+q}^{1/2}(g))$. Now since $\lambda_{\max}(A'A) = \lambda_{\max}(AA')$ for any matrix and since $\Gamma_T^{-1/2}(f_0)\, H \Gamma_{T+q}(g)\, H' \Gamma_T^{-1/2}(f_0) = I$ in view of $y_T = H u_{T+q}$ we obtain the inequality $T^{-1} y_T' \Gamma_T^{-1}(f_0)\, y_T \leqslant T^{-1} u_{T+q}' \Gamma_{T+q}^{-1}(g) u_{T+q}$. Observing that $g \in F_{\mathbb{R}}^{1,2,3}$ and that $(2\pi)^{-1} g$ is the spectral density of $u(t)$ we conclude from Lemma 3.6 in Pötscher [3] that the r.h.s. of the above inequality converges to s. This completes the proof.

REFERENCES

[1] DEISTLER, M., DUNSMUIR, W., AND HANNAN, E. J. (1978). Vector linear time series models: Corrections and extensions. *Adv. in Appl. Probab.* **10** 360–372.
[2] DUNSMUIR, W., AND HANNAN, E. J. (1976). Vector linear time series models. *Adv. in Appl. Probab.* **8** 339–364.
[3] PÖTSCHER, B. M. (1987). Convergence results for maximum likelihood type estimators in multivariable ARMA models. *J. Multivariate Anal.* **21** 29–52.